periodR – an R package to calculate long term cancer survival estimates using period analysis


B Holleczek¹, A Gondos², H Brenner²

¹ Saarland Cancer Registry (Saarbrücken)
² German Cancer Research Center, Division for Clinical Epidemiology and Aging Research (Heidelberg)
periodR – an R package for survival estimation using period analysis

period analysis

- **long term survival** is key outcome parameter reported by cancer registries
- **period analysis**: reflects survival experience within a restricted recent calendar period („left truncation“)
- **method**: cumulative survival rates based on estimated conditional survival for subsequent time intervals after diagnosis (life table methodology)
- up-to-date prediction of results of traditional survival analysis (early detection of changes in survival)
- **empirically evaluated** and widespread use by cancer registries
periodR – an R package for survival estimation using period analysis

period analysis vs. traditional cohort based analysis

<table>
<thead>
<tr>
<th>Diagnose</th>
<th>Nachbeobachtung</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td>1/2 2/3 3/4 4/5 5/6 6/7 7/8 8/9 9/10 10</td>
</tr>
<tr>
<td>1993</td>
<td>1/2 2/3 3/4 4/5 5/6 6/7 7/8 8/9 9/10 10</td>
</tr>
<tr>
<td>1994</td>
<td>1/2 2/3 3/4 4/5 5/6 6/7 7/8 8/9 9/10 10</td>
</tr>
<tr>
<td>1995</td>
<td>1/2 2/3 3/4 4/5 5/6 6/7 7/8 8/9 9/10 10</td>
</tr>
<tr>
<td>1996</td>
<td>1/2 2/3 3/4 4/5 5/6 6/7 7/8 8/9 9/10 10</td>
</tr>
<tr>
<td>1997</td>
<td>1/2 2/3 3/4 4/5 5/6 6/7 7/8 8/9 9/10 10</td>
</tr>
<tr>
<td>1998</td>
<td>1/2 2/3 3/4 4/5 5/6 6/7 7/8 8/9 9/10 10</td>
</tr>
<tr>
<td>1999</td>
<td>1/2 2/3 3/4 4/5 5/6 6/7 7/8 8/9 9/10 10</td>
</tr>
<tr>
<td>2000</td>
<td>1/2 2/3 3/4 4/5 5/6 6/7 7/8 8/9 9/10 10</td>
</tr>
<tr>
<td>2001</td>
<td>1/2 2/3 3/4 4/5 5/6 6/7 7/8 8/9 9/10 10</td>
</tr>
<tr>
<td>2002</td>
<td>1/2 2/3 3/4 4/5 5/6 6/7 7/8 8/9 9/10 10</td>
</tr>
<tr>
<td>2003</td>
<td>1/2 2/3 3/4 4/5 5/6 6/7 7/8 8/9 9/10 10</td>
</tr>
<tr>
<td>2004</td>
<td>1/2 2/3 3/4 4/5 5/6 6/7 7/8 8/9 9/10 10</td>
</tr>
</tbody>
</table>

traditional cohort based estimation of 10 year survival

period analysis based on observed survival between 2002-2004 (left truncation!)

mortality follow up available for end of period (2004)
periodR – an R package for survival estimation using period analysis

is a language and environment for statistical computing and visualizing data

- implements the S language
- available for many OS (Windows, Linux, MacOS)
- freely available under GNU public license
- packages extend R (provide additional functions, datasets, documentation)
- currently available version: R 2.5.1 (released 2007/06/28)
- september 2007: 1139 packages available from CRAN
periodR – an R package for survival estimation using period analysis

periodR implements and extends existing SAS macros for period analysis

implementation/methodological details:

- calculation of **absolute and relative survival** based on conditional survival estimates for **1-year intervals** following diagnosis
- methods for relative survival: **Ederer II** and **Hakulinen**
- **standard errors** (Cutler-Ederer-method),
- **age-standardization** (uses individual weights)
periodR – an R package for survival estimation using period analysis

periodR
the package contains

- functions for **period analysis** (also cohort, complete or hybrid analyses possible)
- functions for data visualization (**plotting** survival curves)
- **datasets** (stomach cancer data, tables with age specific survival probabilities of males and females in Germany 1993-2002)
- integrated **documentation** and executable **examples**
periodR – an R package for survival estimation using period analysis

stomach cancer data
data set of 2302 cancer cases diagnosed in Saarland between 1993 and 2002 (DCO cases already excluded)

variables required for analysis
- sex
- age at diagnosis
- month and year of diagnosis
- month and year of end of follow up
- vital status at end of follow up

for relative survival needed: lifetable data of age and sex specific survival probabilities
periodR – an R package for survival estimation using period analysis

start of R session

> require(periodR)
Lade notiges Paket: periodR
[1] TRUE
>
example analysis 1

> data(stomach)
> data(surv.probs.males)
> data(surv.probs.females)

> result <- period(data=stomach, k=10, surv.probs.males, surv.probs.females, perbeg=2000, perend=2002, method="hakulinen")

> print(result)

estimated survival using period analysis:
Call: period(data=stomach, k=10, surv.probs.males, surv.probs.females, perbeg=2000, perend=2002, method="hakulinen")

method used for relative survival: hakulinen
adjusted for age: no
observations included: 1164
periodR – an R package for survival estimation using period analysis

example analysis 2

<table>
<thead>
<tr>
<th>year</th>
<th>abs.surv</th>
<th>err.abs</th>
<th>rel.surv</th>
<th>err.rel</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.5066</td>
<td>0.0197</td>
<td>0.5287</td>
<td>0.0206</td>
</tr>
<tr>
<td>2</td>
<td>0.3770</td>
<td>0.0191</td>
<td>0.4108</td>
<td>0.0208</td>
</tr>
<tr>
<td>3</td>
<td>0.3128</td>
<td>0.0183</td>
<td>0.3560</td>
<td>0.0209</td>
</tr>
<tr>
<td>4</td>
<td>0.2802</td>
<td>0.0178</td>
<td>0.3344</td>
<td>0.0212</td>
</tr>
<tr>
<td>5</td>
<td>0.2614</td>
<td>0.0175</td>
<td>0.3280</td>
<td>0.0219</td>
</tr>
<tr>
<td>6</td>
<td>0.2338</td>
<td>0.0169</td>
<td>0.3088</td>
<td>0.0224</td>
</tr>
<tr>
<td>7</td>
<td>0.2177</td>
<td>0.0165</td>
<td>0.3029</td>
<td>0.0230</td>
</tr>
<tr>
<td>8</td>
<td>0.2021</td>
<td>0.0161</td>
<td>0.2968</td>
<td>0.0237</td>
</tr>
<tr>
<td>9</td>
<td>0.1831</td>
<td>0.0161</td>
<td>0.2844</td>
<td>0.0251</td>
</tr>
<tr>
<td>10</td>
<td>0.1831</td>
<td>0.0161</td>
<td>0.3015</td>
<td>0.0266</td>
</tr>
</tbody>
</table>

- **year** = follow up year
- **abs.surv** = observed survival at end of follow up year
- **err.abs** = standard error of observed survival
- **rel.surv** = relative survival at end of follow up year
- **err.rel** = standard error of relative survival
periodR – an R package for survival estimation using period analysis

data visualization

plotted survival curves for both sexes for period 2000-2002 (stomach cancer data)
periodR – an R package for survival estimation using period analysis

integrated help
periodR – an R package for survival estimation using period analysis

summary

periodR
• will be freely available under GNU public license
• intended to enhance feasibility and application of period analysis
• next step: model based analysis

further information/software
• http://www.R-project.org (R software, documentation, packages)
• http://www.krebsregister.saarland.de/improve (periodR)
• http://www.imbe.med.uni-erlangen.de/issan/SAS/period (SAS macros for absolute and relative survival rates)
periodR – an R package for survival estimation using period analysis

references

